HHS Public Access

Author manuscript

JAMA Ophthalmol. Author manuscript; available in PMC 2019 September 01.

Development of a Retinopathy of Prematurity Activity Scale and Clinical Outcome Measures for Use in Clinical Trials

Lois E. H. Smith, MD, PhD,
Department of Ophthalmology, Harvard Medical School, Boston Children's Hospital, Boston, Massachusetts

Ann Hellström, MD, PhD,
Department of Clinical Neuroscience at Institute of Neuroscience and Physiology, University of Göteborg, Göteborg, Sweden

Andreas Stahl, MD,
Department of Ophthalmology, University of Freiburg, Freiburg, Germany

Alistair Fielder, MD,
Department of Optometry & Visual Science, City University of London, London, United Kingdom

Corresponding Author: Jonathan M. Davis, MD, The Floating Hospital for Children at Tufts Medical Center, 750 Washington St, Boston, MA 02111, (jdavis@tuftsmedicalcenter.org).

Author Contributions: Drs Davis and Smith had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Smith, Hellström, Stahl, Fielder, Chambers, Moseley, Toth, Wallace, Darlow, Aranda, Hallberg.
Acquisition, analysis, or interpretation of data: Hellström, Stahl, Moseley, Toth, Wallace, Aranda, Davis.
Drafting of the manuscript: Smith, Hellström, Stahl, Fielder, Toth, Darlow, Aranda, Davis.
Critical revision of the manuscript for important intellectual content: Smith, Hellström, Stahl, Chambers, Moseley, Toth, Wallace, Darlow, Aranda, Hallberg, Davis.
Statistical analysis: Hellström.
Administrative, technical, or material support: Stahl, Toth, Aranda, Hallberg, Davis.
Supervision: Smith, Chambers.
Retinopathy of Prematurity Workgroup of the International Neonatal Consortium: Jacob V. Aranda, MD, PhD, State University of New York Downstate Medical Center, Brooklyn, New York; Wiley Chambers, MD, US Food and Drug Administration, Bethesda, Maryland; Brian A. Darlow, MD, University of Otago, Christchurch, New Zealand; Jonathan M. Davis, MD, Tufts Medical Center and the Tufts Clinical and Translational Research Institute, Boston, Massachusetts; Alistair Fielder, MD, City University of London, London, United Kingdom; Dina Apele Freimane, MD, European Medicines Agency, P. Stradins University, Maternal and Child Health Clinics, Riga, Latvia; Boubou Hallberg, MD, Karolinska University Hospital, Stockholm, Sweden; Ann Hellström, MD, PhD, University of Göteborg, Göteborg, Sweden (expert grader); Melissa Liew, MD, Novartis Pharmaceuticals, Basel, Switzerland; Alexandra Mangili, MD, Shire, Zürich, Switzerland; Neil Marlow, MD, Department of Neonatal Medicine, University College London, London, United Kingdom; Jane Moseley, MD, European Medicines Agency, London, United Kingdom; Lois E. H. Smith, MD, PhD, Harvard Medical School, Boston Children's Hospital, Boston, Massachusetts (expert grader); Andreas Stahl, MD, University of Freiburg, Freiburg, Germany (expert grader); Cynthia Toth, MD, Duke University, Durham, North Carolina; and David Wallace, MD, Indiana University, Indianapolis.
Additional Contributions: The following individuals served as the Retinopathy of Prematurity expert graders: Michael Chiang, MD, Department of Medical Informatics and Clinical Epidemiology and Ophthalmology, Casey Eye Institute, Oregon Health & Science University, Alistair Fielder, MD, Department of Optometry & Visual Science, City University of London, London, United Kingdom, Brian Fleck, MD, Department of Ophthalmology, University of Edinburgh, Edinburgh, Scotland, Domenico Lepore, MD, Department of Ophthalmology, Catholic University of the Sacred Heart, Milan, Italy, James Reynolds, MD, Ross Eye Institute, Pediatrics, State University of New York at Buffalo, Buffalo, and David Wallace, MD, Department of Ophthalmology, Indiana University, Indianapolis. Lotta Gränse, Department of Ophthalmology, Lund University, Skåne University Hospital, Lund, Sweden, and Margareta Gebka, Department of Ophthalmology, Lund University, Skåne University Hospital, Lund, Sweden, contributed to the Retinal Photographic Imaging section of the Supplement. Important logistic support was supplied by Lynn Hudson, PhD, Alicia West, and the Critical Path Institute.

Conflict of Interest Disclosures: None reported.

Wiley Chambers, MD,
Division of Transplant and Ophthalmology Products, US Food and Drug Administration, Bethesda, Maryland

Jane Moseley, MD,
Human Medicines Research and Development Support Division, European Medicines Agency, London, United Kingdom

Cynthia Toth, MD,
Department of Ophthalmology, Duke University, Durham, North Carolina

David Wallace, MD,
Department of Ophthalmology, Indiana University, Indianapolis

Brian A. Darlow, MD,
Department of Pathology and Biomedical Science, University of Otago, Christchurch, New Zealand

Jacob V. Aranda, MD, PhD,
Department of Ophthalmology, State University of New York Downstate Medical Center, Brooklyn

Boubou Hallberg, MD, PhD,
Division of Paediatrics, Karolinska University Hospital, Stockholm, Sweden

Jonathan M. Davis, MD,
The Floating Hospital for Children, Tufts Medical Center, Boston, Massachusetts, Department of Pediatrics, Tufts University School of Medicine, Boston, Massachusetts

Retinopathy of Prematurity Workgroup of the International Neonatal Consortium

Abstract

IMPORTANCE—To facilitate drug and device development for neonates, the International Neonatal Consortium brings together key stakeholders, including pharmaceutical companies, practitioners, regulators, funding agencies, scientists, and families, to address the need for objective, standardized clinical trial outcome measurements to fulfill regulatory requirements. Retinopathy of prematurity (ROP) is a disease that affects preterm neonates. The current International Classification of Retinopathy of Prematurity does not take into account all of the characteristics of ROP and does not adequately discriminate small changes in disease after treatment. These factors are critical for evaluating outcomes in clinical trials.

OBSERVATIONS—There is need for an updated ROP acute disease activity and structure scale as well as end-stage structure and ophthalmologic outcome measures designed for use at different ages. The scale and measures, based on current diagnostic methods and treatments, could be used as a guideline for clinical intervention trials. The scale is intended to be validated against retrospective data and revised for use in future trials. An iterative revision process can be accomplished if new measures are added to clinical trials and evaluated at the end of each trial for prognostic value. The new measures would then be incorporated into a new version of the activity scale and the outcome measures revised.

CONCLUSIONS AND RELEVANCE—An ROP activity scale and outcome measures to obtain the most robust and discriminatory data for clinical trials are needed. The scales should be dynamic and modified as knowledge and imaging modalities improve and then validated using data from well-documented clinical trials. This approach is relevant to improving clinical trial data quality.

The International Classification of Retinopathy of Prematurity (ICROP), established in 1984, provided for the first time a standard way to describe retinopathy of prematurity (ROP) that allowed practitioners and centers to evaluate potential interventions.[1] Screening for ROP commenced worldwide, and several multicenter clinical trials evaluating retinal ablation were initiated. With a solid evidence base, blindness from ROP was one of the few causes of childhood visual impairment that could largely be prevented with timely screening and treatment. However, ROP-associated visual disability still occurs in all communities and in epidemic proportions in middle-income countries.[2] The incidence of ROP in some areas is increasing in part because of higher oxygen targets and increased survival of extremely low-gestational-age infants.[3] There is an urgent need to develop new strategies to improve ROP outcomes that can be adopted in all neonatal intensive care unit settings worldwide.

To test new interventions in a clinical setting, there is an unmet need for a revised classification system to accurately document progression (worsening), regression, and recurrence of ROP. The ICROP was developed before anti–vascular endothelial growth factor (VEGF) therapy, which alters the pattern of disease regression and recurrence (eAppendix 1 in the Supplement). The International Neonatal Consortium (INC) consists of key stakeholders in clinical drug development for neonates, including pharmaceutical companies, practitioners, regulators, scientists, and families. The INC is focused on developing objective descriptors that discriminate drug effects at all stages of ROP and outcome measures for clinical trials, which is important to support safety and efficacy claims. Pending the availability of objective outcome measures, uniform implementation of the current classification system and a core set of evaluations will help to facilitate drug development.

At present, intervention studies approach different stages of ROP and can be divided into the following: (1) early prevention (eg, insulinlike growth factor 1, fatty acids, tight control of oxygen delivery following strict target saturations, and neonatology quality improvement programs), (2) prevention of progression (worsening) (eg, propranolol), and (3) late-stage treatment targeting pathologic angiogenesis (eg, anti-VEGF)[4] or the production of angiogenic factors by destruction of the avascular retina (laser photocoagulation)[5] or other pharmacologic interventions (eg, that may promote normal vascularization).

This article addresses 2 challenges. The first challenge is how to collect robust measurements during acute-phase ROP that are sensitive to the response to novel treatments, including reproducible measures of regression, progression, and recurrence. There should be clear standards for objective quantifiable data. The second challenge is how to develop robust short-and long-term outcome measures for clinical trials so that results from an intervention can be reliably ascertained in as short a time as possible but still be indicative of long-term outcome and safety (eAppendix 1 in the Supplement).

Herein, we outline the first iteration of a suggested unbiased approach to an acute disease severity scale derived from the current ICROP classification and based on current diagnostic methods and treatments, as well as clinical outcome measures at different ages. These considerations are intended as a guide for new standards for clinical trials. The acute activity scale and longer-term outcome measures must initially be validated with retrospective data with long-term outcomes and then further modified and revalidated before adoption as a new standard. The scale should evolve with new trials using the standard measurements but also include testing new measurements (such as optical coherence tomography [OCT]), which can then be incorporated (if found to have prognostic value) in a further iteration of the scale.

The ROP clinical trial data quality may be improved with the use of objective quantifiable measures of retinal function and structure in the scales, particularly photographic documentation of areas of nonvascularization, areas of neovascularization, and vessel diameter and tortuosity (plus disease) that cover the entire vascularized retina. High-resolution film-based retinal photography has been replaced by digital photographic technology, which provides prompt access to images for immediate and later evaluation as well as an interface with computer or server storage and electronic transmission. We describe a practical approach to rapidly obtain good-quality fundus photographs (eAppendix 2 in the Supplement). It is possible to obtain objective quantification of plus disease from images using algorithms currently under development only when reliable photographic retinal documentation is available. Further-more, with increased availability of new technologies to evaluate activity and structure (eg, retinal imaging devices based on cell phone cameras, confocal scanning laser ophthalmoscopy, and OCT) (eAppendix 3 in the Supplement) and new quantifiable methods for evaluation of plus disease and extent of retinal vascularization (fluorescein angiography or OCT angiography) (eAppendix 4 in the Supplement), these data can be incorporated into activity and structural scales with appropriate validation.

These considerations are provided in the context of principles of clinical pharmacology and drug development. Rational dose selection derived from pharmacokinetic and pharmacodynamic studies should lead to improved outcomes in future ROP clinical trials.[6] Clinical trials of hypertension in children have found that poor dose selection, lack of consideration of differences between adult and pediatric populations, and lack of pediatric formulations are associated with treatment failure.[7] These factors may also apply to the development of agents for use in ROP. Dose extrapolation from adult ocular drug studies should consider the relative differences in neonatal vitreous volume and gel-like characteristics[8] as well as the retinal surface, which may influence drug distribution and action. Safe and user-friendly ocular drug formulations are urgently needed for neonates and their caregivers. Neonatal animal models of oxygen-induced retinopathy can provide pharmacokinetic and pharmacodynamic data as well as safety and mechanistic data.

Proposal for an ROP Activity and Structure Scale

For clinical trials, it is important to be able to compare the smallest reliably measurable change in ROP activity and structure between visits. For that purpose and to start with an

unbiased initial scale, all possible combinations of the 3 currently used determinants defining ROP severity (zone, stage, and plus disease) were generated. These 22 combinations were sorted by severity based on the judgment of 9 ROP experts (including L.E.H.S., A.H., A.S.) to produce a first-generation acute ROP activity and structure scale. There was no attempt at this time to create a scale with equal intervals of severity. Using the proposed acute ordinal ROP severity scale, it should be possible to determine (1) the rate of ROP progression during and after treatment and (2) the stability of this improvement over time (ie, how long it takes until a relapse in ROP occurs). To best capture progression, regression, and recurrence of ROP, a second iteration of the scale would likely include higher-resolution descriptors. For that purpose, all trials aimed at drug development should collect zone, stage, pre-plus and plus disease, and extent of stage 3 disease by quadrants as well as other possible quantifiable measures, such as OCT or fluorescein angiography. The new indicators would then be evaluated at the completion of the trial(s) for their prognostic value and incorporated into a new iteration of the scale if found to be valid.

The ROP activity and structure scale described here (first iteration) needs to be validated by demonstrating that different grades in the scale are associated with different risks of important long-term outcomes using retrospective and prospective clinical trial data and modified as more measurable indicators become available (such as OCT and quantifiable measurements of vessel width, plus disease, and extent of nonvascularized area and neovascular growth). Although OCT currently is limited to evaluation of the posterior part of the fundus (equivalent to zone 1), it can potentially be useful for evaluating the location of the foveal center, aggressive posterior ROP,[9] and plus disease[10] and distinguishing between stages 4a and 4b in eyes with retinal detachment.[9] eAppendix 3 in the Supplement gives an overview of OCT for structural efficacy end points. Further iterations of the scale would incorporate new indicators if they are proved to determine outcome.

ROP Activity and Structure Scale

In this first iteration of an acute ROP activity and structure scale, the overall mean of the rankings, zones, and stages that require treatment according to the Swedish National Quality Register (SWEDROP) study[11] are listed in Table 1. All of these grades align on the lower (more severe) end of the scale. There are areas of ambiguity with higher SDs where the experts show less agreement (Table 1). However, there are clear cuts in the severity scale, marked by large differences between one level of severity and the next (ie, high Δ values). Two high Δ (distance to the next-higher overall score) values were used to distinguish 3 subcategories of ROP: mild, moderate, and severe (Figure).

Structuring ROP severity into an integrated scale has strengths and weaknesses. Strengths include the overall consensus among grading experts at the mild and severe ends of the spectrum and the ability to identify 3 subcategories of mild, moderate, and severe ROP. In addition, all ROP stages that require treatment according to current guidelines are at the severe end of the scale. The severity scale's limitations are visible in the upper mild and moderate spectrum of ROP, where graders differed considerably in their individual scores. Although some of these theoretical combinations may not exist clinically, they are included in the first iteration of the scale for completeness. To study the clinical prevalence of each

component of the proposed scales, a member of the steering board of the Swedish National Quality Register (A.H.) evaluated data on eye, zone, stage, and plus disease from 3248 preterm infants examined from 2008 to 2015 in Sweden and categorized them according to the scale components. The scale rankings 4, 11, 12, and 15 were never seen in this cohort. Scale rankings 10, 16, and 17 were seen in only 1 of 3248 eyes. Thus, lack of clinical relevance of some of the scale stages may have added to the ambiguity seen in the expert gradings (eg, scale 4 indicates ROP zone III stage 1 with plus disease). However, to avoid prematurely excluding any possible combination that might (at least theoretically) arise in a clinical trial, all possible combinations were included. After further analysis using retrospective clinical trial data, rankings may change; then a revalidated scale could be used in future clinical trials. Further refinement of the scale by adding refined clinical descriptors, such as pre-plus disease and quadrants as well as weighting components of the scale rankings to account for components differentially affecting ROP outcome, may be helpful. However, this refinement must also be validated with high-quality retrospective and prospective clinical trial data.

Long-term Ophthalmic Outcome Measures

In addition to capturing acute ROP activity and structure changes during a clinical trial, it is essential to document ROP treatment outcomes in a standardized way(Table 2). For structural end points, ROP outcome measurements are different from acute-phase ROP measurements. Although activity and structure of acute ROP should be recorded throughout a clinical trial, the structural end-stage ROP outcome is usually measured only once, most often as the primary or key secondary end point. Ideally, different ROP trials should use the same structural end points to allow for comparison of trials even when the timing of study visits, interventions, or primary end point analyses differ. The time of the end point analysis may vary from study to study. However, it is reasonable to evaluate structural outcomes of ROP at 6 months (26 weeks) and 66 weeks corrected gestational age (CGA). In addition to structural and morphologic end points, functional outcomes should be captured as they become measurable with increasing age and cooperation of the infant.

The proposed timing should be rational and well justified in each trial. Discussions with regulators should focus on the selection of the primary outcome of the trial, which may need to be tailored to the pharmacologic properties and intended therapeutic use of the medicinal product. This approach could include the absence of active ROP or any unfavorable structural retinal outcome according to the definitions below.

Subgroup Investigations

In addition to the above standard examinations, further investigations that require specialized equipment and expertise may yield important additional information in clinical trials, such as OCT (eAppendix 3 in the Supplement), fluorescein angiography (with oral fluorescein dye and confocal scanning laser ophthalmoscopy in an outpatient setting or intravenous angiography with the patient under anesthesia) (eAppendix 4 in the Supplement), and electrophysiology (electroretinogram).

Follow-up for ROP Prevention and Treatment Trials

There are many compelling reasons to systematically follow up very premature infants; a substantial proportion will have ongoing health issues or neurodevelopmental delays that are important for parents and neonatal intensive care unit teams to know about. Such follow-up is especially important because future problems can be prevented or ameliorated with early intervention if identified early enough. Current and future therapies to prevent or treat ROP may have systemic and local effects. For instance, surgical therapies that are undertaken with a general anesthetic may affect the central nervous system, and intravitreal anti-VEGF therapy is known to be systemically absorbed and may potentially affect other organs.

The methods and timing of follow-up will vary according to the specific objectives. In comparative effectiveness research and trials of new therapies, safety will be a major consideration and will be guided by the drug class profile evaluation and pharmacodynamics. Many trials of neonatal therapies have an early primary outcome and are not powered to detect mortality or longer-term secondary outcomes. In addition, the adverse effects of therapies can mimic the common morbidities seen with prematurity.[13] Therefore, large numbers of infants and randomized treatment assignment may be required to provide evidence of harm or safety.

Neurodevelopmental assessment is typically performed at 18 to 24 months CGA, with the later time points better correlating with future outcomes. The limitations of such an assessment as an indicator of future cognitive and functional outcomes have been discussed in numerous publications.[14–16] Composite outcomes that combine cognition, motor, and sensory measures may not describe children adequately; for instance, a child who has isolated severe deafness or visual impairment may achieve what society judges as normal functionality with the provision of appropriate aids.

The current criterion standard for assessment of neurodevelopmental status at 2 years CGA is the Bayley Scales of Infant and Toddler Development–Third Edition. However, issues with normative values mean that local population norms or contemporaneous term-born control groups are ideally required for interpretation of test results. As the child becomes older, tools that assess day-to-day functioning, executive function, visual processing and coordination, and fine motor integration become more feasible. Validated parental questionnaires (eg, Parent Report of Children's Abilities–Revised for Preterm Infants) may be a useful option to determine moderate or severe disabilities but generally not milder or more subtle problems.

Imperative for Drug Development to Prevent and Treat ROP

The global epidemic of ROP and the expected increase in visually impaired or blind children worldwide underscore the urgency of pharmacologic strategies, drug development, and clinical trials to prevent or treat ROP.[17] Determination of robust outcome measures for clinical trials that are acceptable to practicing health care professionals, the scientific community, pharmaceutical companies, and regulatory agencies is crucial. Graded severity scales must be validated with high-quality clinical trial data. The scoring systems described

above are based on ICROP criteria stage, zone, and plus disease, which have been proven to be clinically relevant. Thus, reliability of scales based on this proven approach should be ensured. An improved grading system, including pre-plus and plus disease, quadrants of ROP, and OCT-based indicators, could be incorporated when validated. Retinal findings should be documented with complete retinal images that can be evaluated and reexamined as needed. Image documentation will enhance agreement among investigators on the grading of the primary outcomes because there is currently wide variability in the classification of plus disease by ROP experts.[18,19]

Long-term follow-up studies that include neurobehavioral, visual (Table 2), and cognitiveoutcomesand relevant clinical outcomesshould be evaluated in subsequent safety protocols. Harmonization of protocols, outcome measures, and other key elements of the trials as was achieved in the oxygen saturation trials will enhance the ability to combine data across trials.[3] This task should facilitate further data analyses and improvement of subsequent trial designs.

Limitations of the ROP Activity Scale

The scale described in this communication has limitations. Some of the 22 options in the scale do not exist clinically. They were included in the first unbiased iteration, but components of the scale should be removed (or collapsed) if there is no prognostic value after validation of the scale with a retrospective study cohort with available long-term outcome measures. A 22-component scale is likely to be too complex for clinical use. A 3-level scale is likely to be too simple for disease classification. The scale in this iteration is not linear. It will be difficult to develop and evaluate a linear scale. It will take many iterations of a scale over time with ongoing validation of new components (such as OCT measures) to create a construct that is more useful than the current SWEDROP construct. It may be difficult to convince trialists to incorporate new measures that will not be used in their trial design but used for validation for future trials and will incur costs that will only benefit the community long term.

Conclusions

In this Special Communication, we suggest a first iteration of an acute activity scale and longer-term outcome measures to be used to obtain the most robust data from ROP clinical trials and to propose more refined descriptors for acute-phase ROP (which must be validated for prognostic value). This first scale has limitations. Some of the 22 options in the scale do not exist clinically. We anticipate that with evaluation of the current version of the scale using retrospective trial data (that also has long-term outcome measures), many levels of the scale will be found to add no prognostic information and therefore will be eliminated or collapsed. This first scale is a starting point for future iterations that can incorporate quantitative data. We hope that this scale can be a data-driven, unbiased approach to constructing and modifying a scale that should continue to evolve as more quantifiable measures of ROP become widely available and tested for prognostic value and are added to further iterations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding/Support: The International Neonatal Consortium (INC) is supported in part by grant U18FD005320-01 from the US Food and Drug Administration to the Critical Path Institute and through annual dues of member companies and by grant EY017017 from the National Eye Institute (Dr Smith) and European Commission FP7 project 305485 PREVENT-ROP (Dr Smith) for manuscript preparation.

Role of the Funder/Sponsor: The support from the INC includes preparation, review, and approval of the manuscript and decision to submit the manuscript for publication.

REFERENCES

1. International Committee for the Classification of Retinopathy of Prematurity. The International Classification of Retinopathy of Prematurity revisited. Arch Ophthalmol 2005;123(7):991–999. doi: 10.1001/archopht.123.7.991 [PubMed: 16009843]
2. Blencowe H, Lawn JE, Vazquez T, Fielder A, Gilbert C. Preterm-associated visual impairment and estimates of retinopathy of prematurity at regional and global levels for 2010. Pediatr Res 2013;74(suppl 1):35–49. doi:10.1038/pr.2013.205 [PubMed: 24366462]
3. Askie LM, Darlow BA, Finer N, et al.; Neonatal Oxygenation Prospective Meta-analysis (NeOProM) Collaboration. Association between oxygen saturation targeting and death or disability in extremely preterm infants in the Neonatal Oxygenation Prospective Meta-analysis Collaboration. JAMA 2018;319(21):2190–2201. doi:10.1001/jama.2018.5725 [PubMed: 29872859]
4. VanderVeen DK, Melia M, Yang MB, Hutchinson AK, Wilson LB, Lambert SR. Anti-vascular endothelial growth factor therapy for primary treatment of type 1 retinopathy of prematurity: a report by the American Academy of Ophthalmology. Ophthalmology 2017;124(5): 619–633. doi: 10.1016/j.ophtha.2016.12.025 [PubMed: 28341474]
5. Early Treatment for Retinopathy of Prematurity Cooperative Group. Revised indications for the treatment of retinopathy of prematurity: results of the Early Treatment for Retinopathy of Prematurity randomized trial. Arch Ophthalmol 2003;121(12): 1684–1694. doi:10.1001/archopht. 121.12.1684 [PubMed: 14662586]
6. Wallace DK, Kraker RT, Freedman SF, et al.; Pediatric Eye Disease Investigator Group (PEDIG). Assessment of lower doses of intravitreous bevacizumab for retinopathy of prematurity: a phase 1 dosing study. JAMA Ophthalmol 2017;135 (6):654–656. doi:10.1001/jamaophthalmol.2017.1055 [PubMed: 28448664]
7. Benjamin DK Jr, Smith PB, Jadhav P, et al. Pediatric antihypertensive trial failures: analysis of end points and dose range. Hypertension 2008;51 (4):834–840. doi:10.1161/HYPERTENSIONAHA. 107.108886 [PubMed: 18332283]
8. Le Goff MM, Bishop PN. Adult vitreous structure and postnatal changes. Eye (Lond) 2008;22(10): 1214–1222. doi:10.1038/eye.2008.21 [PubMed: 18309340]
9. Chavala SH, Farsiu S, Maldonado R, Wallace DK, Freedman SF, Toth CA. Insights into advanced retinopathy of prematurity using handheld spectral domain optical coherence tomography imaging. Ophthalmology 2009;116(12):2448–2456. doi:10.1016/j.ophtha.2009.06.003 [PubMed: 19766317]
10. Maldonado RS, Yuan E, Tran-Viet D, et al. Three-dimensional assessment of vascular and perivascular characteristics in subjects with retinopathy of prematurity. Ophthalmology 2014; 121(6):1289–1296. doi:10.1016/j.ophtha.2013.12.004 [PubMed: 24461542]
11. ClinicalTrials.gov. The Early Treatment for Retinopathy of Prematurity Study (ETROP) NCT00027222 https://clinicaltrials.gov/ct2/show/NCT00027222 Accessed January 11, 2018.
12. Lundgren P, Hård A-L, Wilde Å, Löfqvist C, Smith LEH, Hellström A. Implementing higher oxygen saturation targets reduced the impact of poor weight gain as a predictor for retinopathy of prematurity. Acta Paediatr 2018;107(5):767–773. doi:10.1111/apa.14049 [PubMed: 28872709]

13. Le Cras TD, Markham NE, Tuder RM, Voelkel NF, Abman SH. Treatment of newborn rats with a VEGF receptor inhibitor causes pulmonary hypertension and abnormal lung structure. Am J Physiol Lung Cell Mol Physiol 2002;283(3):L555–L562. doi:10.1152/ajplung.00408.2001 [PubMed: 12169575]

14. Rysavy MA, Marlow N, Doyle LW, et al. Reporting outcomes of extremely preterm births. Pediatrics 2016;138(3):e20160689. doi:10.1542/peds.2016-0689 [PubMed: 27516525]

15. Kilbride HW, Aylward GP, Doyle LW, Singer LT, Lantos J. Prognostic neurodevelopmental testing of preterm infants: do we need to change the paradigm? J Perinatol 2017;37(5):475–479. doi: 10.1038/jp.2017.12 [PubMed: 28252658]

16. Johnson S, Marlow N. Early and long-term outcome of infants born extremely preterm. Arch Dis Child 2017;102(1):97–102. doi:10.1136/archdischild-2015-309581 [PubMed: 27512082]

17. Quinn GE. Retinopathy of prematurity blindness worldwide: phenotypes in the third epidemic. Eye Brain 2016;8:31–36. doi:10.2147/EB.S94436 [PubMed: 28539799]

18. Campbell JP, Kalpathy-Cramer J, Erdogmus D, et al.; Imaging and Informatics in Retinopathy of Prematurity Research Consortium. Plus disease in retinopathy of prematurity: a continuous spectrum of vascular abnormality as a basis of diagnostic variability. Ophthalmology 2016;123(11):2338–2344. doi:10.1016/j.ophtha.2016.07.026 [PubMed: 27591053]

19. Stahl A, Krohne TU, Eter N, et al.; Comparing Alternative Ranibizumab Dosages for Safety and Efficacy in Retinopathy of Prematurity (CARE-ROP) Study Group. Comparing Alternative Ranibizumab Dosages for Safety and Efficacy in Retinopathy of Prematurity: a randomized clinical trial. JAMA Pediatr 2018;172(3):278–286. doi:10.1001/jamapediatrics.2017.4838 [PubMed: 29309486]

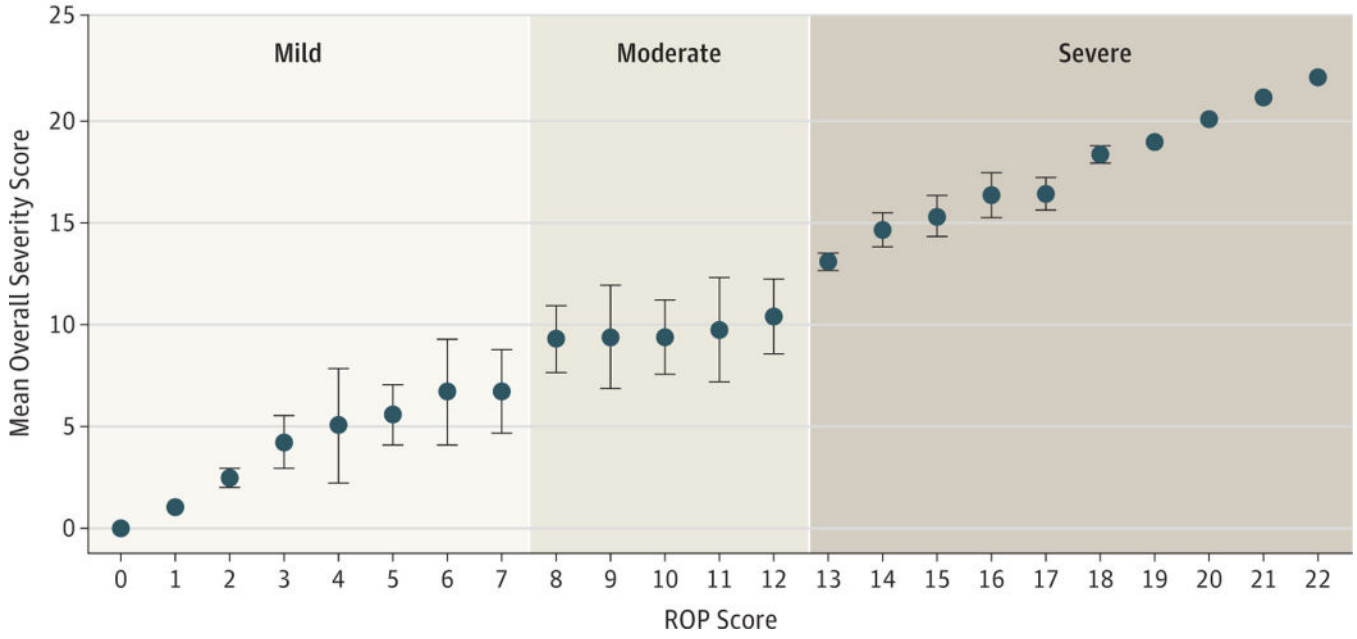

Figure. Retinopathy of Prematurity (ROP) Activity and Structure Scores vs Mean Expert Gradings

Error bars indicate SDs. Short error bars indicate good consensus with mild and severe ROP; large error bars indicate ambiguity in moderate ROP stages.

Table 1.

ROP Activity and Structure Score Evaluation by ROP Experts

	Score by Expert No.											
Scale/Zone/Stage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Overall Score**	**Category**	**Δ (SD)[a]**
0/None/none	0	0	0	0	0	0	0	0	0	0.0		1.0 (0.00)
1/III/1	1	1	1	1	1	1	1	1	1	1.0		1.4 (0.00)
2/III/2	3	3	2	2	3	2	2	2	3	2.4		1.8 (0.50)
3/II/1	5	4	7[b]	3	5	3	3	3	5	4.2		0.8 (1.31)
4/III/1+	2[b]	10[b]	4	9[b]	2[b]	6	5	5	2[b]	5.0	Mild	0.6 (2.79)
5/III/3	7	5	3[b]	5	6	5	4	8[b]	7	5.6		1.1 (1.50)
6/III/2+	4[b]	11[b]	5	10[b]	4[b]	7	6	9[b]	4[b]	6.7[c]		0.0 (2.58)
7/II/2	9[b]	6	8	4[b]	8	4[b]	8	4[b]	9[b]	6.7[c]		2.6 (2.05)
8/II/3	10	7[b]	9	6[b]	9	11	10	11	10	9.2[c]		0.1 (1.62)
9/III/3+	8	12[b]	6[b]	12[b]	7[b]	8	9	14[b]	8	9.3[c]		0.0 (2.54)
10/I/1	11	8	11	7[b]	10	9	11	6[b]	11	9.3[c]	Moderate	0.3 (1.83)
11/II/1+	6[b]	13[b]	10	11	12[b]	12[b]	7[b]	10	6[b]	9.7[c]		0.7 (2.54)
12/I/	12	9	12	8[b]	11	10	12	7[b]	12	10.3		2.7 (1.83)
13/II/2+	13	14	13	13	13	13	13	12	13	13.0		1.6 (0.47)
14/II/3+	14	15	14	16	14	14	14	16	14	14.6		0.7 (0.83)
15/I/1+	15	16	16	14	16	16	16	13	15	15.2		1.0 (1.03)
16/I/3	17	18	15	17	15	15	15	17	17	16.2[c]		0.1 (1.13)
17/I/2+	16	17	17	15	17	17	17	15	16	16.3[c]	Severe	1.9 (0.82)
18/I/3+	18	19	18	18	18	18	18	19	18	18.2		0.7 (0.42)
19/Any zone/AP-ROP	19	19	19	19	19	19	19	18	19	18.9		1.1 (0.31)
20/Any zone/4a	20	20	20	20	20	20	20	20	20	20.0		1.0 (0.00)
21/Any zone/4b	21	21	21	21	21	21	21	21	21	21.0		1.0 (0.00)

	Score by Expert No.											
Scale/Zone/Stage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Overall Score**	**Category**	**Δ (SD)[a]**
22/Any zone/5	22	22	22	22	22	22	22	22	22	22.0		0.0 (0.00)

Abbreviations: AP, aggressive posterior; ROP, retinopathy of prematurity.

[a] Δ Indicates the distance to the next-higher overall score. SD is the deviation between the 9 graders.

[b] Greater than 2-point deviation from overall score.

[c] Almost identical overall scores of neighboring score.

Table 2.

ROP Ophthalmologic Outcome Evaluation

	Time of Evaluation		
Outcome	**66 wk CGA**	**2.5 y**	**5 y**
Vision (with best correction)			
Light reaction	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND
Fix and follow 5-cm toy	Yes/no (5 cm)	Yes/no (at 30 cm)	Yes/no (at 30 cm)
Ability to identify single optotypes, 0.4 LEA test at 3-m distance (test passed if 3 of 4 symbols are correctly identified after prescribing glasses if indicated according to recommendations by the AAO, and retest vision after 3 mo)[11,12]	NA	Value; prescription given	Value; prescription given
ETDRS letters (test distance at4m [1 or 0.5m if poor vision] right eye, left eye, and both eyes after prescribed glasses if indicated according to recommendations by the AAO, and retest vision after 3 mo)[11,12]	NA	NA	Value right eye, left eye, both eyes; prescription given
Stereoacuity if present: record in arc seconds (TNO, Lang, other)	NA	NA	Outcome (method)
Intraocular pressure method	NA	Value (method)	Value (method)
Cycloplegic refraction, sphere/cylinder/axis; method (retinoscopy/autorefractor/other)	Value (method)	Value (method)	Value (method)
Ocular palsy	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND
Manifest strabismus (ET, XT, HT)	ET, XT, HT/none/U/ND	ET, XT, HT/none/U/ND	ET, XT, HT/none/U/ND
Nystagmus in primary position	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND
Retina			
Normal attached retina to periphery	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND
Normal central retina, peripheral retina scarred, quiescent (seen after successful laser treatment)	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND
Normal central retina, peripheral retina not vascularized, a visible border band may be present in the midperiphery (see below), can be seen after anti-VEGF therapy	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND
Fibrovascular organization (seen in x quadrants), can include posterior retinal folds, foveal dragging, vitreoretinal traction, elevated preretinal structures, dragging of retinal vessels or retrolental membrane obscuring the posterior pole; in addition, retinoschisis (on OCT if available)	Yes/no/U/ND (if yes, No. of quadrants)	Yes/no/U/ND	Yes/no/U/ND
Dense vitreous or preretinal hemorrhage (with poor visualization of the fundus)	Yes/no/U/ND	NA	NA
ROP stages 4 or 5; if yes, then subcategories (4a, 4b, or 5) should be recorded	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND
Disease recurrence, including presence of a new stage 2 or 3 disease or leakage on fluorescein angiography (if available) with or without plus disease and with or without new retinal hemorrhages	Yes/no/U/ND	NA	NA
Optic nerve			
Optic nerve hypoplasia	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND
Optic nerve atrophy	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND

Outcome	Time of Evaluation		
	66 wk CGA	2.5 y	5 y
Large excavation of the optic disc (>1/2 disc area)/ other optic nerve anomaly (describe)	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND
Lens			
Normal clear lens	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND
Cataract; if yes, describe localization and density.	Yes/no/U/ND	Yes/no/U/ND	Yes/no/U/ND
Contrast sensitivity, outcome, and method	NA	NA	Outcome (method)
Visual fields outcome and method (Humphrey, other)[3]	NA	NA	Outcome (method)
Additional treatment for recurrent ROP, ocular and systemic adverse events	NA	NA	NA

Abbreviations: AAO, American Academy of Ophthalmology; CGA, corrected gestational age; ET, esotropia; ETDRS, Early Treatment for Diabetic Retinopathy Study; HT, hypertropia; LEA, pediatric vision testing; NA, not applicable; ND, not done; OCT, optical coherence tomography; ROP, retinopathy of prematurity; TNO, measure of stereopsis; U, unable to assess; VEGF, vascular endothelial growth factor; XT, extotropia.